# ON THE NOTION OF RANDOM INTENSITY

A.I. Yashin

March 1985
WP-85-013



INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
2361 Laxenburg, Austria

# ON THE NOTION OF RANDOM INTENSITY

A.I. Yashin

## INTRODUCTION

The problems of explaining the observed trends in mortality, morbidity and other kinds of individuals' transitions generated the numerous attempts of incorporating the covariates into the survival models. First models use the deterministic constant factors as explanatory variables [1, 2, 3]. Gradually it became clear that the random and dynamic nature of the covariates should also be taken into account [4, 5, 6]. This understanding has led to the fact that the notion of random intensity became widely used in the analysis of the asymptotic properties of the maximum likelihood and Cox-regression estimators [7, 8].

Having the clear intuitive sense the notion of random intensity can be introduced in different ways. The traditional way is to define the intensity in terms of probability distributions of the failure time [9, 10, 11]. Another way appeals to the martingale theory [12] and defines the intensity in terms of the predictable process, called "compensator" [13, 14]. For the deterministic rates and simple cases of stochastic intensities there are already results that establish a one-to-one correspondence between two definitions. The correspondence is reached by the probabilistic representation results for compensator [13, 14]. Martingale theory guarantees the existence of the predictable compensator in more general cases. However the results on the probabilistic representation similar to simple cases are still unknown. Meanwhile such representation is crucial, for instance, in the analysis of the relations between the duration of the life cycle of some unit and stochastically changing influential variables. This paper shows the result of such representation for some particular case. The generalization on the more general situations is straightforward. The consideration will use some basic notions of a "general theory of processes" [12, 15].

## PRELIMINARIES

Let $T$ be the stopping time defined on some probability space $(\Omega, H, \mathbf{H}, P)$, where $\mathbf{H} = (H_t)_{t \geq 0}$ is the right-continuous nondecreasing family of $\sigma$-algebras in $\Omega$, such that $H_\infty = H$ and $H_0$ is completed by $P$-zero sets from $H$. If distribution of $T$ is absolutely continuous the traditional definition of the intensity $\lambda(t)$, related to the stopping time $T$ is as follows

$$\lambda(t) = \frac{\frac{d}{dt} P(T \leq t)}{P(T > t)} \quad . \tag{1}$$

Martingale characterization defines the rate in terms of the process $A(t)$ which is supposed to be $\mathbf{H}$-predictable and such that the process $M(t)$ defined as

$$M(t) = I(T \leq t) - A(t)$$

is an $\mathbf{H}$-adapted martingale. It turns out that for process $M(t)$ to be martingale, $A(t)$ should have the form

$$A(t) = \int_0^t I(u \le T)\lambda(u)du \quad .$$

where $\lambda(u)$ is given by (1). The family **H** in this case is generated by the indicator process $X_t = I(t \ge T)$. If the stopping time $T$ is correlated with some random variable $z(\omega)$, then the traditional approach defines the intensity in terms of conditional probabilities

$$\lambda(t,z) = \frac{\frac{d}{dt}P(T \le t \mid z)}{P(T > t \mid z)} \quad . \tag{2}$$

Martingale characterization shows that the process $M^z(t)$

$$M^z(t) = I(T \le t) - \int_0^t I(u \le T)\lambda(u,z)du$$

is a martingale with respect to the family of $\sigma$-algebras $\mathbf{H}^z$, generated in $\Omega$ by the indicator process $I(t \ge T)$ and random variable $z$. In the case of a discontinuous conditional distribution function for $T$, formula (2) should be corrected. The notion of cumulated intensity $\Lambda(t,z)$ is more appropriate in this case. The formula for it is

$$\Lambda(t,z) = \int_0^t \frac{dP(T \le u \mid z)}{P(T \ge u \mid z)} \tag{3}$$

and martingale characterization is respectively [13, 14]:

$$M^z(t) = I(T \le t) - \int_0^t I(u \le T)d\Lambda(u,z) \quad . \tag{4}$$

The situation is not so clear however if one has the random processes, say $Y_t$, correlated with stopping time $T$. Assume for instance that $Y_t$ simulates the changes of the physiological variables of some patient in hospital and $T$ is the time of death. It is clear that in this case the process $Y_t$ should terminate at time $T$, so observing $Y_t$ one can tell about the alive/dead state of the patient, that is, can observe the death time.

It is clear also that the state or the whole history of the physiological variable influences the chances of occurring death. The question is how can one specify the random intensity in terms of conditional probabilities in order to establish the correspondence between intuitive traditional and martingale definitions of the random intensities.

The idea to use $\Lambda(t,Y)$ in the form

$$d\Lambda(t,Y) = \frac{dP(T \le t \mid H_t^y)}{P(T \ge t \mid H_t^y)}$$

where $\sigma$-algebra $H_t^y$ is generated in $\Omega$ by the process $Y_u$ up to time $t$ fails because $H_t^y$ contains the event $\{T \le t\}$. Taking $H_{t-}^y$ instead of $H_t^y$ seems to improve the situation, however the event $\{T \ge t\}$ is measurable with respect to $H_{t-}^y$.

So to find the proper formula for random intensity one needs to get the probabilistic representation result for $H^y$-predictable compensator.

**RESULT FORMULATION**

We will demonstrate the result and the ideas of proof on a simple particular case. The generalization on a more general situation is straightforward.

Assume that stopping time $T(\omega)$ and the Wiener process $W_t(\omega)$ are defined on some probability space $(\Omega, H, P)$ and are independent for any $t \ge 0$. Define

$$Y_t = W_t - I(T \le t)\, W_t \quad .$$

Let $\mathbf{H}^y = (H_t^y)_{t \ge 0}$, $\mathbf{H} = (H_t^w)_{t \ge 0}$ where

$$H_t^y = \sigma\{Y_u, u \le t\}$$

$$H_t^w = \sigma\{w_u, u \le t\} \quad .$$

The following statement is true.

**Theorem.** $\mathbf{H}^y$*-predictable compensator* $A^y(t)$ *of the process* $I(T \le t)$ *have the following representation*

$$A^y(t) = \int_0^t I(u \le T) \frac{d_u P(T \le u \mid H_u^w)}{P(T \ge u \mid H_u^w)} \quad . \tag{5}$$

The proof of this theorem will be followed by some auxilliary lemmas.

Let $H_t^x = \sigma\{X_u, u \le t\}$, where $X_t = I(T \le t)$. It is clear that $H_t^y = H_t^w \vee H_t^x$. Denote $H^x = H_\infty^x = H_T$.

The next assertion underlines the important property of the stopping times with respect to $\mathbf{H}^y$.

**Lemma** 1. *For any* $\mathbf{H}^y$ *stopping time* $\tau$ *one can find the random variables* $S$ *and* $S'$ *such that indicator* $I(S \ge s)$ *is* $H_s^w$ *measurable for any* $s \ge 0$*; indicator* $I(S' \ge s)$ *is* $H_T$ *measurable for any* $s \ge 0$ *and:*

$$I(\tau \ge s) I(s \le T) = I(S \ge s) I(s \le T)$$

$$I(\tau \ge s) I(T < s < \infty) = I(S' \ge s) I(T < s < \infty) \quad .$$

Let us first prove the first equality. From the definition of $H_s^x$ it follows that

$$H_{s-}^x \cap \{s \le T\} = \{s \le T\} \quad ;$$

therefore

$$(H_s^w \vee H_{s-}^x) \cap \{s \le T\} = H_s^w \cap \{s \le T\}$$

or by virtue of coincidence of the sets $H_t^y \cap \{t \le T\}$ and $(H_t^w \vee H_t^x) \cap \{t \le T\}$ and continuity of the flow of $\sigma$-algebras $\mathbf{H}^w$:

$$H_{s-}^y \cap \{s \le T\} = H_s^w \cap \{s \le T\} \quad .$$

Taking the set $\{\tau < s\}$ from $\mathbf{H}_{s-}^y$ one can find the set $D_s$ from $H_s^w$ such that

$$\{\tau < s\} \cap \{s \le T\} = D_s \cap \{s \le T\} \quad .$$

Note now that for $r < s$

$$\{D_r \cup D_s\} \cap \{s \le T\} = D_s \cap \{s \le T\} \quad .$$

In fact

$$D_r \cap \{s \le T\} = D_r \cap \{r < T\} \cap \{s \le T\} = \{\tau < r\} \cap \{r \le T\} \cap \{s \le T\} =$$

$$\{\tau < r\} \cap \{\tau < s\} \cap \{r \le T\} \cap \{s \le T\} \subseteq \{\tau < s\} \cap \{s \le T\} = D_s \cap \{s \le T\} \quad .$$

Define the random variable $S$ by the relations

$$\{S < s\} = \bigcup_{r<s} D_r ,\ s \ge 0 \quad .$$

Where $r$ and $s$ are rational numbers we have

$$\{\tau < s\} \cap \{s \le T\} = \{S < s\} \cap \{s \le T\}$$

or

$$\{\tau \ge s\} \cap \{s \le T\} = \{S \ge s\} \cap \{s \le T\}$$

which completes the proof of the first equality. The second equality can be proved in a similar way.

The next lemma deals with the representation of $\mathbf{H}^w$-adapted square integrable martingales as stochastic integrals with respect to Wiener process $W_t$.

**Lemma 2.** *Let $M = (M_t)_{t \ge 0}$ be a $\mathbf{H}^w$-adapted square integrable martingale. Then there is an $\mathbf{H}^w$-adapted process $f(s,\omega)$, $s \ge 0$ such that for any $t$*

$$E \int_0^t f^2(s,\omega) ds < \infty$$

*and*

$$M_t = M_0 + \int_0^t f(s,\omega) dW_s \quad .$$

The proof of this lemma may be found in [13].

The next statement will play an important role in the characterization of $\mathbf{H}^w$-predictable processes.

**Lemma 3.** *Any* $\mathbf{H}^w$*-well measurable process is* $\mathbf{H}^w$*-predictable.*

*Proof.* Let $\tau$ be the arbitrary stopping time with respect to $\mathbf{H}^w$. Denote by $\Lambda(t)$ $\mathbf{H}^w$-predictable compensator of the process $I(\tau \le t)$. According to the definition of compensator the process $M_t = I(\tau \le t) - \Lambda(t)$ is the bounded martingale with respect to $\mathbf{H}^w$. According to Lemma 2 there exists $\mathbf{H}^w$-adapted function $f(s,\omega)$ such that

$$M_t = M_0 + \int_0^t f(s,\omega) dW_s \quad .$$

By virtue of the continuity of the stochastic integral's trajectories, the right-hand side of this equality is $\mathbf{H}^w$-predictable. This fact yields the $\mathbf{H}^w$-predictability of the indicator $I(\tau \le t)$ for any $t$ that in turn yields the $\mathbf{H}^w$-predictability of the stopping time $\tau$. The predictability property for any stopping time with respect to $\mathbf{H}^w$ yields the coincidence of the well-measurable and predictable $\sigma$-algebras.

**Lemma 4.** $\mathbf{H}^y$*-adapted process* $Z = (Z_t)_{t\ge0}$ *is* $\mathbf{H}^y$*-predictable if and only if there is an* $\mathbf{H}^w$*-measurable process* $X = (X_t)_{t\ge0}$ *and* $\mathbf{H}^w_T$*-well measurable process* $X' = (X'_t)_{t\ge0}$ *such that*

$$X_t I(t \le T) = Z_t I(t \le T) \tag{6}$$

$$X'_t I(T < t) = Z_t I(T < t) \quad . \tag{7}$$

*Proof.* **Necessity**. Let $Z_t = I(t \le \tau)$ where $\tau$ is an arbitrary stopping time with respect to $\mathbf{H}^y$. According to the result of Lemma 1 we have:

$$I(t \le \tau) I(t \le T) = I(t \le S) I(t \le T)$$

where $S$ is some random variable, such that indicator $I(t \le S)$ is $H_t^w$-measurable for any $t \ge 0$. Setting $Y_t = I(t \le S)$ we get the necessity of the condition (6) for the process $Z_t = I(t \le \tau)$.

Denote by $L$ the set of bounded $\mathbf{H}^y$-adapted functions which satisfy relation (6). It is easy to see that $L$ is closed with respect to monotonic convergence to bounded functions and contains the set of indicators $I(t \le \tau)$, where $\tau$ is arbitrary stopping times with respect to $\mathbf{H}^y$. This set of indicators is the algebra; it contains the 1, and generates the $\mathbf{H}^y$-predictable $\sigma$-algebra in $\Omega \times R_+$. The necessity of the condition (6) of the lemma follows as a result of montone classes theorem [15]. In the same way the necessity of condition (7) can be proved.

**Sufficiency**. Note that for any arbitrary $\mathbf{H}^w$-adapted and left continuous process $l_t$ the process $I(t \le T)l_t$ is $\mathbf{H}^y$-adapted and left continuous; therefore

$$P(\mathbf{H}^w) \cap ]|0,T]| \subset P(\mathbf{H}^y)$$

$$P(\mathbf{H}^w) \cap ]|T,\infty[| \subset P(\mathbf{H}^y)$$

where $P(\mathbf{H}^w)$ and $P(\mathbf{H}^y)$ denote $\mathbf{H}^w$-predictable and $\mathbf{H}^y$-predictable $\sigma$-algebras, respectively. According to Lemma 3 the process $Y_t$ from the relation (6) is $\mathbf{H}^w$-predictable, and process $X$ from (7) is $\mathbf{H}^w$-predictable. Consequently the processes $X_t I(t \le T)$ and $X_t' I(T \le t)$ are $\mathbf{H}^y$-predictable.

Define the process $\Lambda(t)$ by the equality

$$\Lambda(t) = \int_0^{t \wedge T} \frac{d_s P(T \le s \mid H_s^w)}{P(T \ge s \mid H_s^w)}$$

The following statement concerns the measurability property of $\Lambda(t)$ which is important for predictable specification of the compensator.

**Lemma 5.** *The process $\Lambda(t)$ is $\mathbf{H}^y$-predictable.*

*Proof.* For $\Lambda(t)$ one can write

$$\Lambda(t) = \int_0^{t\wedge T} \frac{d_s P(T \le s \mid H_s^w)}{P(T \ge s \mid H_s^w)} = I(t \le T)\int_0^t d_s P(T \le s \mid H_s^w) + I(t > T)\int_0^T \frac{d_s P(T \le s \mid H_s^w)}{P(T \ge s \mid H_s^w)} \quad .$$

The first addendum on the right-hand side is an $\mathbf{H}^y$-predictable process in accordance with Lemma 4. The second is an $\mathbf{H}^y$-adapted, left-continuous process and consequently it is also $\mathbf{H}^y$-predictable.

The next lemma demonstrates the last effort in proving the main result.

**Lemma 6.** *The process $Y_t = I(T \le t) - \Lambda(t)$ is $\mathbf{H}^y$-martingale.*

*Proof.* Consider the process

$$M(t) = I(t \ge T) - \Lambda(t) \quad .$$

It is clear that $E|M(t)| < \infty$. Let us prove that

$$E(M(t) \mid H_s^y) = M(s) \quad .$$

We have

$$E(M(t) \mid H_s^y) = E(I(t \ge T) \mid H_s^y) - E(\Lambda(t) \mid H_s^y) \quad .$$

The first term can be transformed into

$$E(I(t \ge T) \mid H_s^y) = I(s \ge T) + I(s < T)\,\frac{E(I(s < T)I(T \le t) \mid H_s^w)}{P(T > s \mid H_s^w)} \quad . \tag{8}$$

In order to transform the second addendum let us assume for simplicity that the conditional distribution of $T$ is absolutely continuous and consequently $\Lambda(t)$ can be represented in the form

$$\Lambda(t) = \int_0^t I(u \le T)\lambda(w,u)du$$

where $\lambda(w,u)$ is $\mathbf{H}^w$-adapted process

$$\lambda(w,u) = \frac{\frac{dP}{du}(T \le u \mid H_u^w)}{P(T \ge u \mid H_u^w)} \tag{9}$$

we have

$$E(\int_0^t I(u \le T)\lambda(w,u)du \mid H_s^y) =$$

$$= \int_0^s I(u < T)\lambda(w,u)du + \frac{I(s < T)E(\int_s^t I(u \le T)\lambda(w,u)du \mid H_s^w)}{P(T > s \mid H_s^w)} \ . \tag{10}$$

Taking into account formula (9) for $\lambda(\omega,u)$, the last addendum can be transformed into the form

$$\frac{I(s < T)}{P(T > s \mid H_s^w)} E(\int_s^t P(u < T \mid H_u^w)\lambda(w,u)du \mid H_s^w) =$$

$$= \frac{[P(T \le t \mid H_s^w) - P(T \le s \mid H_s^w)]I(s < T)}{P(T > s \mid H_s^w)}$$

$$= \frac{I(s < T)P(s < T \le t \mid H_s^w)}{P(T > s \mid H_s^w)} \ . \tag{11}$$

Subtracting (10) from (8) and taking into account (11), one can see that

$$E(M(t) \mid H_s^y) = M(s)$$

The uniqueness of $\mathbf{H}^y$-predictable compensator yields the coincidence of $\Lambda(t)$ and $A^y(t)$ that completes the proof.

**Remark.** It turns out that even if the $\sigma$-algebra $H_t^y$ has the more general structure than in the example above, for instance, $H_t^y = H_t^z \vee H_t^w$, the result of the probabilistic representation of the $H^y$-predictable compensator of the process

$X(t) = I(T \leq t)$ given by formula (5) is true.

**Example**. The relevance of the results such as formula (5) becomes evident from the following example. Assume that Wiener process $W_t$ and stopping time $T$ are inter-related and the random intensity of occurrence $T$ is $\lambda(t)W_t^2$. Formula (5) gives immediately the form of the conditional survival function

$$S(t \mid W_0^t) = P(T > t \mid H_t^2) = e^{-\int_0^t \lambda(u) W_u^2 du}$$

In survival analysis the stopping time $T$ is associated with the death or failure time and the research is often focused on the properties of the survival function $S(t) = P(T > t)$ [8, 11, 10, 7]. The straightforward way of its calculation is the averaging of the conditional survival function $S(t \mid W_0^t)$. It turns out that (see for instance [16])

$$E e^{-\int_0^t \lambda(u) W_u^2 du} = e^{-\int_0^t \lambda(u) E(W_u^2 \mid T > u) du}$$

For the Wiener process the conditional mathematical expectation on the right-hand side of this formula can be easily calculated (the condition $W_0 = 0$ is used there)

$$E(W_t^2 \mid T > t) = \gamma(t)$$

where $\gamma(t)$ is the solution of the differential equation

$$\dot{\gamma}(t) = 1 - 2\lambda(t)\gamma^2(t) \quad , \quad \gamma(0) = 0 \quad .$$

When $\lambda(t)$ is constantly equal, say, to $\frac{1}{2}$ the straightforward calculations lead to the formula

$$E e^{-\frac{1}{2}\int_0^t W_u^2 du} = \frac{1}{\sqrt{cht}}$$

which coincides with the result based on the Cameron and Martin formula [13].

## CONCLUSION

Formula (5) can be generalized in more complex cases including the sequence of stopping times and semimartingale as an influential stochastic process. It can be useful in the field of survival analysis, reliability theory and risk analysis. It shows which particular conditional distribution functions should be used in specification of the random intensities. The specification of the influential process and the measurement schemas provide the particular forms for the distributions and the intensities. Some examples, related to the biomedical and demographical applications are discussed in [17, 18, 16].

## REFERENCES

1. D.R. Cox, "Regression Models and Life Tables," *Journal of Royal Statistical Society* **B 34**, pp.187-220 (1972).

2. D.R. Cox, "Partial Likelihood," *Biometrika* **A 62** , pp.269-276 (1975).

3. K. Bailey, "The Asymptotic Joint Distribution of Regression and Survival Parameter Estimates in the Cox Regression Model," *The Annals of Statistics* **11**(1), pp.39-48 (1983).

4. P.K. Anderson and R. Gill, *Cox Regression Model for Counting Processes: A Large Sample Study*, Statistical Research Unit, Danish Medical Research Council, Danish Social Science Research Council, Copenhagen (1981). RR-81-6.

5. R.D. Gill, "Understanding Cox's Regression Model: A Martingale Approach," *Journal of American Statistical Association* **79**(386), pp.441-447, June 1984 (1984).

6. R.L. Prentice and S.G. Self, *Asymptotic Distribution Theory for Cox-type Regression Models with General Relative Risk Form*, 1983.

7. D.R. Cox and D. Oakes, *Analysis of Survival Data*, Chapman and Hall, London (1984).

8. R.C. Elandt-Johnson and N.L. Johnson, *Survival Models and Data Analysis*, John Wiley and Sons, New York (1979).

9. R.E. Barlow and F. Proschan, *Statistical Theory of Reliability and Life Testing*, Holt, Rinehart and Winston, Inc., New York (1975).

10. J.F. Lawless, *Statistical Models and Methods for Lifetime Data*, John Wiley and Sons, New York (1982).

11. W. Nelson, *Applied Life Data Analysis* , John Wiley and Sons, New York (1982).

12. J. Jacod, *Calcules Stochastique et Probleme de Martingales. Lecture Notes in Mathematics, Vol. 714*, Springer-Verlag, Heidelberg (1979).

13. R.S. Liptzer and A.N. Shiryaev, *Statistics of Random Processes*, Springer-Verlag, Berlin and New York (1978).

14. J. Jacod, "Multivariate Point Processes: Predictable Projection, Radon-Nicodim Derivatives, Representation of Martingales," *Zeitschrift für Wahrscheinlichkeitstheorie und Verw. Gebiete* **31**, pp.235-253 (1975).

15. C. Dellacherie, *Capacities et Processus Stochastiques (Capacities and Stochastic Processes)*, Springer-Verlag, Berlin and New York (1972).

16. A.I. Yashin, *Dynamics in Survival Analysis: Conditional Gaussian Property Versus Cameron-Martin Formula.* *WP-84-107*, International Institute for Applied Systems Analysis, Laxenburg, Austria (1984).

17. A.I. Yashin, K.G. Manton, and J.W. Vaupel, *Mortality and Aging in a Heterogeneous Population: A Stochastic Process Model with Observed and Unobserved Variables.* *WP-83-81*, International Institute for Applied Systems Analysis, Laxenburg, Austria (1983).

18. A.I. Yashin and K.G. Manton, *Evaluating the Effects of Observed and Unobserved Diffusion Processes in Survival Analysis of Longitudinal Data.* *CP-84-58*, International Institute for Applied Systems Analysis, Laxenburg, Austria (1984).